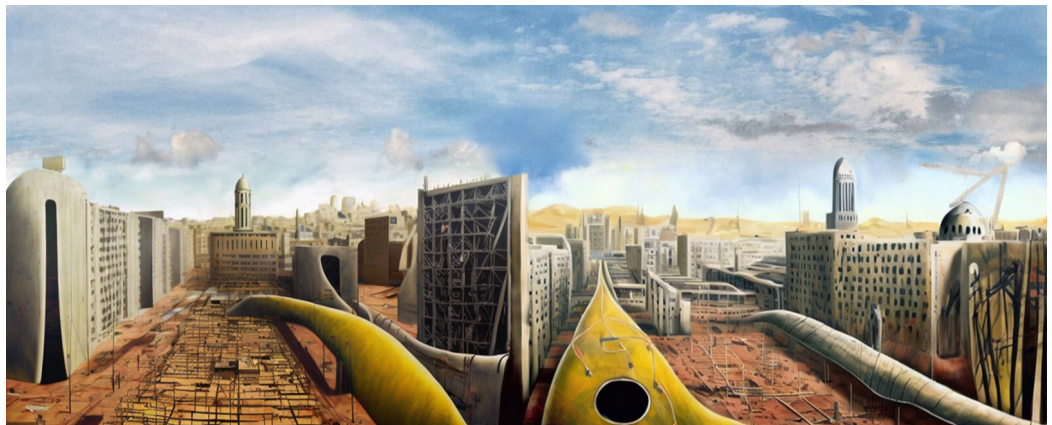
44° Convegno Internazionale
dei Docenti delle Discipline della Rappresentazione
Congresso della Unione Italiana per il Disegno

Transizioni d'immagini e architetture al tempo dell'IA. Modelli semantici in cerca di autore

Maurizio Unali
Giovanni Caffio
Fabio Zollo

Abstract

Il saggio sintetizza l'esito di un lungo percorso di ricerca incentrato sullo studio del ruolo rappresentativo e conformativo del modello semantico in architettura, attualizzato rispetto alle più recenti innovazioni tecnoculturali introdotte dal concetto di Intelligenza Artificiale. L'idea di modello semantico è stata verificata e sperimentata attraverso i più recenti processi elaborativi della rappresentazione che in vario modo utilizzano l'IA, dalla tecnologia del *text-to-image* alla modellazione parametrica morfo-generativa (e oltre). Si tratta di progettualità post-digitali che consentono nuove transizioni di senso alimentate dalle ibridazioni culturali e dai riciclaggi estetici riscontrabili in ogni ambito disciplinare, ma che nella rappresentazione architettonica trovano un proficuo laboratorio di sperimentazione per configurare le forme dell'abitare virtuale, anche nella dimensione del metaverso. La ricerca è stata strutturata in due fasi. Nella prima fase, dopo aver studiato i principali antefatti e riferimenti concettuali e visivi, sono state catalogate e classificate alcune fra le molteplici progettualità elaborate sul tema da architetti, designer e artisti, evidenziando semantiche, metodi e tecniche di elaborazione. Nella seconda fase sono state predisposte delle progettualità tematiche che, fra ricerca e didattica, hanno consentito di elaborare 'transizioni' di immagini e spazi attraverso tecnologie riferibili in generale all'IA. Infine sono state pubblicate le immagini elaborate e illustrate le metodologie adottate.

Parole chiave
Intelligenza Artificiale, modelli semantici, text-to-image, rappresentativo/conformativo, abitare virtuale

Fantasy. *La città analoga nell'epoca della IA*, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Arduino Cantafora, Città Analoga, Costa, Gianni Rodari, Grammatica della Fantasia'.

doi.org/10.3280/oa-1016-c464

Introduzione

La ricerca presentata è l'ultimo esito di un lungo percorso di studio (tuttora in corso) incentrato sull'elaborazione del ruolo rappresentativo e conformativo [Unali 2018] del disegno per il 'Modello Semantico' in architettura e nel design [Unali, Caffio 2023], qui attualizzato rispetto alle più recenti innovazioni tecnoculturali introdotte soprattutto dal concetto di Intelligenza Artificiale, con le conseguenti progettualità interdisciplinari.
La ricerca è stata strutturata in due fasi elaborative, qui sintetizzate in tre capitoli [1].
Nella prima fase – 'Rilevamento dello stato dell'arte' – dopo aver studiato i principali antefatti e riferimenti concettuali e visivi che l'argomento attinge dalla storia della rappresentazione (e non solo), sono state catalogate e classificate alcune fra le molteplici progettualità elaborate sul tema da architetti, designer e artisti, evidenziando semantiche, metodi e tecniche di elaborazione.
Nella seconda fase della ricerca – 'Sperimentazioni' –, dopo delle prime prove applicative prevalentemente di ordine estetico (modelli semantici) e tecnologico (metodi e tecniche) – alimentate soprattutto dalle conoscenze acquisite nella precedente fase –, sono state predisposte delle progettualità tematiche rispetto a una struttura più matura che, fra ricerca e didattica, ci ha consentito di elaborare consapevoli 'transizioni' di immagini e spazi attraverso tecnologie riferibili in generale all'IA, affrontando così i vari aspetti aperti dall'argomento.

Modelli semantici in cerca di autore, fra '*Soft Machine*' e '*Hard Machine*'

Sin dagli esordi della rivoluzione informatica, ogni linguaggio espressivo si è interrogato sui complessi e affascinanti rapporti fra la nostra umana '*Soft Machine*' – citando il titolo del romanzo di William S. Burroughs – e quella artificiale '*Hard Machine*' del digitale (come *HAL 9000* in *2001: Odissea nello spazio* di Stanley Kubrick). Gli esempi sono infiniti, dalla letteratura al cinema, dalla scienza all'arte, fino alla filosofia e oltre.
Nella prima fase della ricerca svolta – 'Rilevamento dello stato dell'arte' – lo studio dei riferimenti che hanno in più modi alimentato il pensiero e l'azione del progetto di architettura nella dimensione creativa oggi attribuibile ai processi dell'IA [Sacchi 2022], sono stati fondamentali e numerosi.
Fra questi, citiamo qui, per motivi di spazio, solo pochissimi antefatti e riferimenti concettuali per ricordare quel prezioso filo rosso contenuto nella storia della rappresentazione: *L'opera d'arte nell'epoca della sua riproducibilità tecnica* (1936) di Walter Benjamin; *The Medium is the Message*, anticipato dal '67 da Marshall McLuhan; i *Quadri Specchianti* virtuali di Michelangelo Pistoletto (dagli anni '60); la sintesi concettuale proposta da Peter Eisenman per la Virtual House (1997), fra i primi esempi di un uso conformativo-creativo dei processi di modellazione; il concetto di *Big Data* studiato nel '99 dai ricercatori Bryson, Kenwright, Cox, Ellsworth e Haimes [Sacchi 2021] *AI, Artificial Intelligence* (2001) di Steven Spielberg, che apre l'argomento al grande pubblico; ChatGPT (*Chat Generative Pre-trained Transformer*), un *chatbot* basato sull'intelligenza artificiale creato da OpenAI nel 2022; la vendita all'asta da Christie's nel 2018 del *Ritratto di Edmond Belamy,* un'opera d'arte generata con l'IA dal Collettivo Obvious. Sulle conoscenze acquisite si fonda la seconda fase della ricerca – Sperimentazioni –, dove sono state realizzate rappresentazioni tematiche attraverso l'uso dell'IA (dal *text-to-image* alla modellazione morfo-generativa) rispetto a una struttura di lavoro che ci ha consentito di elaborare consapevoli 'transizioni' di immagini e di spazi.
Le rappresentazioni elaborate sono state poi strutturate rispetto a due principali linee di ricerca individuate nelle nostre sperimentazioni tecnoculturali dell'IA: transizioni semantiche di immagini piane; transizioni di conformazioni spaziali tridimensionali. All'interno del primo gruppo, utilizzando soprattutto la piattaforma Midjourney, sono stati progettati quattro macrotemi di sperimentazione: 1. Esercizi di Stile (figg. 1-3); 2. Ibridazioni architettoniche (figg. 4-6); 3. Morfemi, coppie duali e coppie oppositive (figg. 7-9); 4. Fantasia (immagine di copertina). Nel secondo gruppo, dopo l'esperienza acquisita dai precedenti processi elaborativi, sono state sperimentate delle trasformazioni morfologiche di spazi in 3D (fig. 10).
Dopo l'analisi dei risultati conformativi raggiunti dalla ricerca, si conferma il grande potenziale teorico (semantico) e visivo (tecnoculturale) che l'IA oggi offre alla ricerca e alla didattica della rappre-

sentazione. Creatività condivise fra *Soft* e *Hard Machine* (semplificando molto), che generano modelli semantici 'in cerca di autore' – in quanto 'progettazione partecipata' –, esito di un nuovo processo elaborativo in cui confluiscono e si intersecano (nella consapevolezza dei riferimenti storici) processi culturali, idee e immagini propedeutiche al progetto di nuove forme dell'abitare. Progettualità che consentono nuove transizioni di senso alimentate dalle ibridazioni culturali e dai riciclaggi estetici riscontrabili in ogni ambito disciplinare, ma che nella rappresentazione architettonica trovano un proficuo laboratorio di sperimentazione per configurare le forme dell'abitare virtuale contemporaneo [Caffio, Unali 2022], anche nella dimensione attuale del metaverso [2] postdigitale [Unali 2019].

Fig. 1. Tav. 1A (1-5). Esercizi di Stile, monotematici autoriali, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Le Corbusier Architecture'.

Fig. 2. Tav. 1B (1-5). Esercizi di Stile, monotematici autoriali, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Louis Kahn Architecture'.

Fig. 3. Tav. 1C (1-5). Esercizi di Stile, monotematici autoriali, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Aldo Rossi Architecture'.

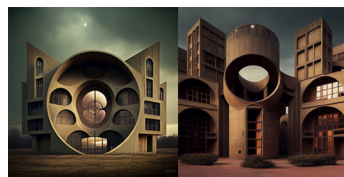



Fig. 4. Tav. 2A (1-5). Ibridazioni architettoniche, autoriali e tipologiche, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Le Corbusier Architecture, Imaginary Prisons Giovanni Battista Piranesi'.

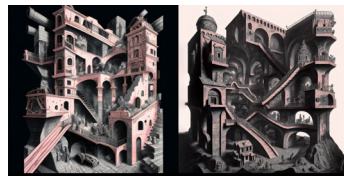
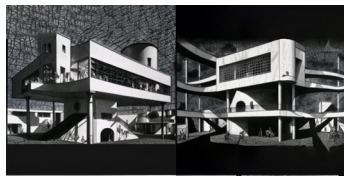



Fig. 5. Tav. 2B (1-5). Ibridazioni architettoniche, autoriali e tipologiche, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Imaginary Prisons Giovanni Battista Piranesi, La Muralla Roja by Bofil, painting Escher style'.

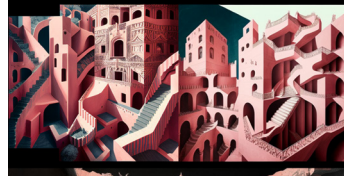
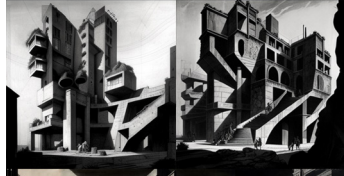



Fig. 6. Tav. 3B (1-5). Ibridazioni architettoniche, autoriali e tipologiche, 2023. Immagine generata con Midjourney, prompt degli Autori, testo: 'Louis Kahn, Aldo Rossi Teatro del Mondo'.

Dal riconoscimento di immagini al *text-to-image* fino al *text-to-3D*

La maggior parte degli esperimenti attualmente in atto nell'ambito delle diverse IA disponibili per studiosi e neofiti dell'architettura, alle varie scale del progetto – dall'oggetto all'edificio fino alla città –, lavorano nel campo della creazione di immagini bidimensionali che alludono a tridimensionalità architettoniche a prima vista riconoscibili e credibili ma che celano spesso

incoerenze proiettive e/o costruttive latenti. L'elaborazione di questo variegato immaginario architettonico nasce soprattutto dalla traduzione di *prompt* testuali in immagini, una sorta di traslitterazione di parole-chiave che innescano interpolazioni di *pixel* fino alla graduale solidificazione in un'immagine grazie alla selezione e intervento di un operatore umano/designer. È importante ricordare che questo prolifico processo di creazione di figurazioni sintetiche è, in parte, il portato di una tecnologia nata per scopi diversi, ovvero il riconoscimento di immagini o visione artificiale, una branca del più ampio settore dell'apprendimento automatico.

Fig. 7. Tav. 3A (1-5). Morfemi e Coppie Duali, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo, text: 'Franco Purini, Morfemi, Axonometric Projections'.

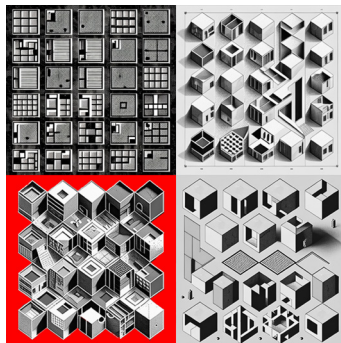



Fig. 8. Tav. 3B (1-5). Morfemi e Coppie Duali, 2023. Immagine generata con Midjourney, *prompt* degli Autori, testo: 'Concept of Full and Empty in Architecture'.

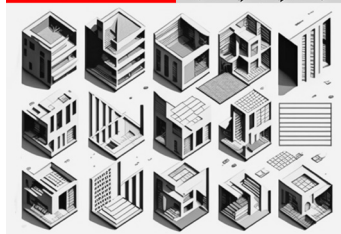
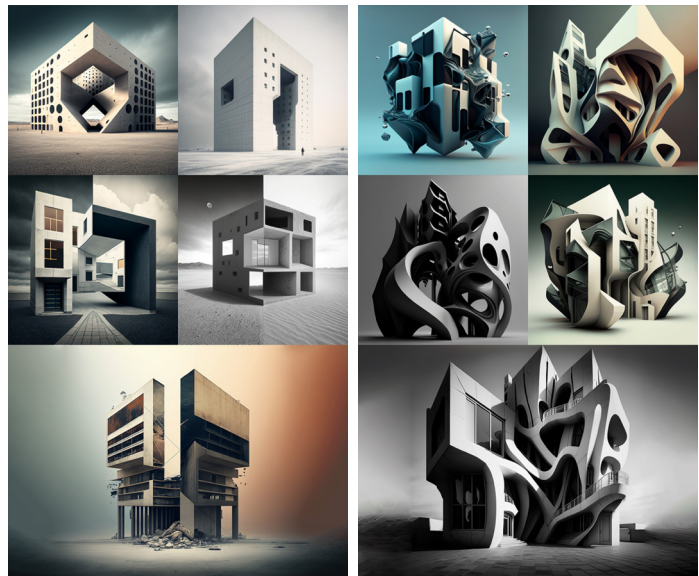



Fig. 9. Tav. 3C (1-5). Morfemi e Coppie Duali, 2023. Immagine generata con Midjourney, prompt degli Autori, testo: 'Fluid Architecture combinated with Rigid (Solid) Architecture'.

È il 2015 quando l'ingegnere di Google Alexander Mordvintsev scopre che una rete neurale, opportunamente addestrata, può essere in grado non solo di riconoscere e classificare delle immagini ma può rovesciare il processo creando nuove immagini a partire da quelle iniziali, di fatto aprendo la strada a un uso creativo e 'artistico' della IA. Nelle parole di Mordvintsev: "L'esperimento consisteva nel provare a creare una rete che aggiungesse dettagli a un'immagine reale per effettuare un aumento della sua risoluzione. È successo che ha aggiunto alcuni dettagli, ma non quelli che mi aspettavo" [3].
Questa scoperta, apparentemente casuale, ha condotto allo sviluppo di DeepDream [Mordvintsev, Olah, Tyka 2015], il primo software in grado di creare immagini dal sapore 'artistico', caratterizzate da un tono onirico e allucinato che hanno incuriosito e generato ampi dibattiti all'interno della critica artistica [Rayner 2016]. Queste nuove possibilità sono state poi ulteriormente sviluppate arrivando alla capacità di trasferire caratteristiche stilistiche o morfologiche da un contesto a un altro o a implementare particolari abilità nell'analisi e riproduzione di selezionati set di immagini come possono essere le opere prodotte da un singolo artista [Gatys, Ecker, Bethge 2015].
Tuttavia, il punto di svolta più significativo è legato alla invenzione delle reti antagoniste generative (GAN - *Generative Adversial Network*) a opera dell'informatico Ian GoodFellow [Goodfellow et al. 2014]. Queste reti neurali si basano su una coppia di reti che operano in competizione una con l'altra. Una ha il compito di generare le immagini, l'altra quello di confrontarle rispetto a un set di dati di riferimento e formazione. Si sviluppa, così, una vera e propria dialettica interna, fatta di migliaia di prove ed errori, tra un sistema che produce instancabilmente immagini e un altro che le scarta fino a che non corrispondono ai criteri di giudizio prestabiliti. Questo processo ha lo scopo di addestrare la rete generatrice di immagini e, una volta raggiunto, mentre la rete discriminante si ferma, quella generatrice continua a produrre nuove rappresentazioni. Esistono diverse forme di GAN. Una delle più comuni è la StyleGAN [Karras, Laine, Aila 2018] che usa il procedimento dell'interpolazione per generare nuove variazioni a partire da un repertorio di migliaia di immagini. Altre reti GAN, come CycleGAN [Zhu, Park, Isola, Efros 2017], invece di un singolo set di dati, utilizzano due set

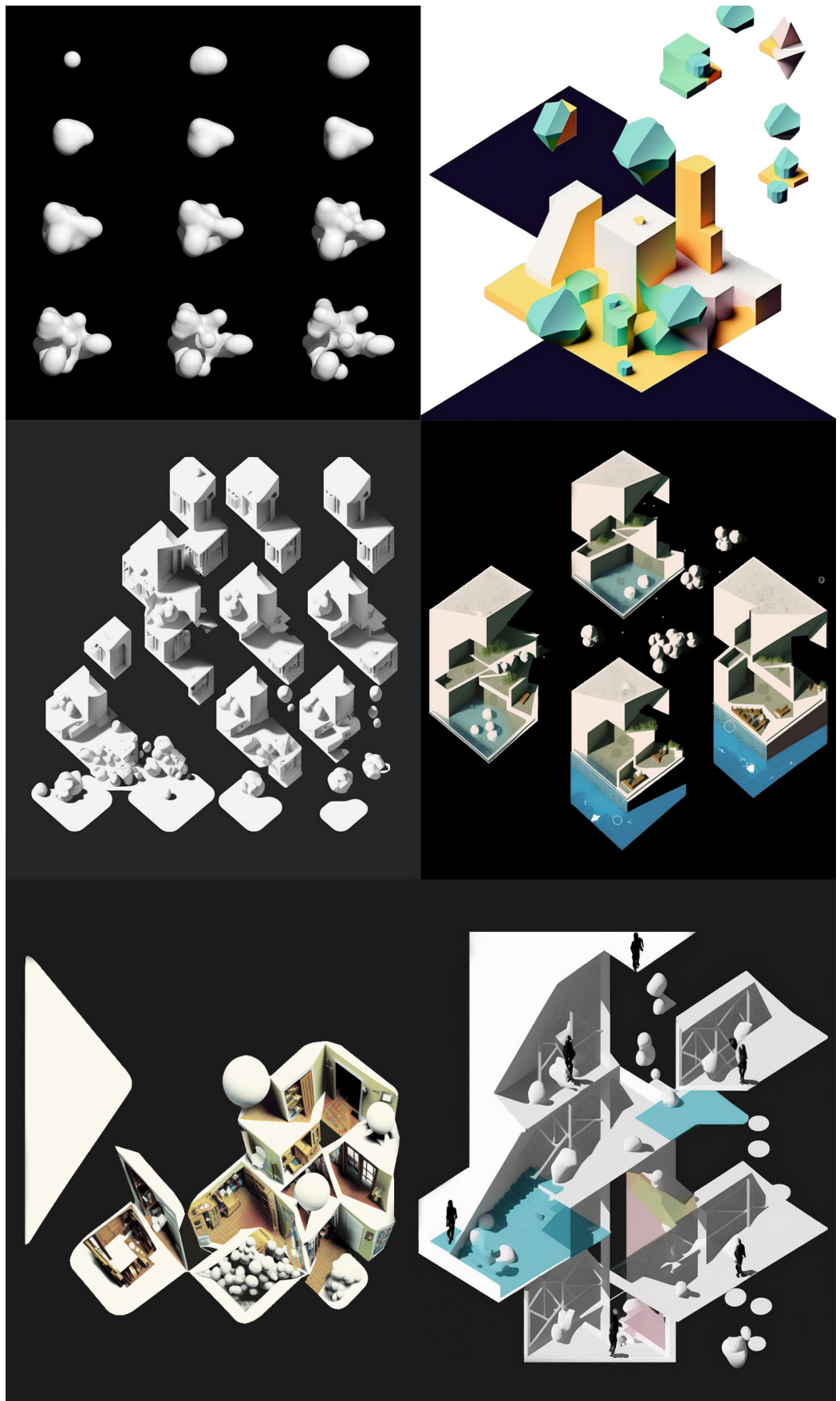



Fig. 10. Tav. 5 (1-6). Transizioni di conformazioni spaziali tridimensionali, 2023. Tav. 5.1. Modellazione morfo-generativa. Tavv. 5.2.-5.6. Interazioni fra modellatore e text- to-image. Immagine generata con Midjourney, prompt degli Autori, testo: 'Axonometric Projections, Transitions'.

dicotomici lavorando non più attraverso l'interpolazione bensì per estrapolazione. In questo modo le immagini prodotte appaiono il risultato di un processo più 'creativo' in quanto introducono elementi diversi e innovativi rispetto a quelli ricavabili esclusivamente da un unico set.
Saltando qualche passaggio, risulta interessante pensare che tali processi elaborativi sono sempre stati presenti nell'essere umano che, attraverso il digitale amplia le sue potenzialità immaginative. A tale riguardo ricordiamo l'esperienza nei primi anni '70 di Gianni Rodari che introduceva il suo metodo del 'binomio fantastico': "Una storia può nascere solo da un 'binomio fantastico' [...]. Occorre una certa distanza tra le due parole, occorre che l'una sia sufficientemente estranea all'altra, e il loro accostamento discretamente insolito, perché l'immaginazione sia costretta a mettersi in moto per istituire tra loro una parentela, per costruire un insieme (fantastico) in cui i due elementi estranei possano convivere" [Rodari 2010, p. 16].
Tornando alla recente storia delle IA, le innovazioni sono poi continuate con la nascita di CLIP (*Contrastive Language–Image Pre-training*) [4], un sistema per la generazione di immagini che impiega GPT-3 (*Generative Pre-trained Transformer 3*), un metodo di elaborazione di linguaggio naturale che sfrutta le potenzialità del *deep learning* per creare testo apparentemente prodotto dall'uomo ed è anche in grado di usare i contenuti presenti in rete per associare stringhe di testo esplicativo alle immagini.
Le potenzialità del CLIP diventano poi esponenziali quando abbinate a VQGans (*Vector Quantized Generative Adversarial Network*) [Esser, Rombach, Ommer 2021] una rete neurale in grado di combinare reti neurali convoluzionali per generare immagini ad alta risoluzione. Il risultato di questa unione è la possibilità di generare immagini bidimensionali a partire da *prompt* testuali come avviene con gli ormai noti Dall-e 2 [5], Midjourney [6] e Stable Diffusion [7].
A differenza della generazione di figure bidimensionali, i cui risultati sono ogni giorno che passa più efficaci e realistici, attualmente sono ancora in corso di sperimentazione sistemi in grado, con la stessa raffinatezza, di tradurre *prompt* testuali in forme tridimensionali anche se, grazie all'aumento della potenza computazionale, al miglioramento degli algoritmi, al maggior numero di concorrenti e ai crescenti investimenti, non è difficile immaginare future sostanziali innovazioni.
Tra i progetti che lavorano nella direzione della generazione di strutture tridimensionali possiamo citare Point-e e DreamFusion. Il primo è un sistema di apprendimento automatico sviluppato da OpenAI (la stessa società che sviluppa Dall-e 2 e ChatGPT) che genera oggetti tridimensionali sulla base di *prompt* testuali. In realtà Point-e non crea una forma conclusa bensì nuvole di punti nello spazio che un secondo sistema di IA converte in mesh [Wiggers 2022].
DreamFusion [Poole, Jain, Barron, Mildenhall 2022], invece, utilizza un modello di diffusione *text-to image 2D* pre-addestrato per realizzare una sintesi *text-to 3D*. Il modello di diffusione 2D funge da base iniziale per l'ottimizzazione di un generatore di immagini parametriche. In una procedura simile a quella usata in DeepDream, il modello tridimensionale è generato inizialmente in modo casuale (un *Neural Radiance Field*, o NeRF) in modo che le sue rappresentazioni 2D da angoli casuali appaiano sufficientemente coerenti. Il modello 3D risultante può essere poi visualizzato da qualsiasi angolazione, illuminato da sorgenti luminose o inserito in uno spazio 3D.
Al di là dei progressi computazionali ancora da venire, qui ci preme evidenziare come siano molti gli architetti e i ricercatori che già sperimentano diversi approcci e soluzioni all'uso delle IA per la generazione di nuove creatività architettoniche. Partiamo citando le fondamentali ricerche condotte a Matias Del Campo attraverso i suoi testi [Del Campo 2022] e le sperimentazioni architettoniche condotte con il suo studio SPAN (con Sandra Manninger) utilizzando una rete neurale convoluzionale standard (CNN), chiamata rete neurale convoluzionale grafica (GCNN), per modellare ed eseguire l'intero processo di progettazione a modelli e informazioni tridimensionali [Del Campo, Leach 2022, p. 45].
Lo studio Habidatum combina reti neurali GAN e intervento umano, principi estetici e processi di *form-finding*, per generare le configurazioni della serie Endless Skyscraper [8]. Esquivel, Bugni e Jaminet sperimentano le reti GAN per ricercare reinterpretazioni 3D del codice serliano nel progetto *The Serlio Code* [9]. Ricordiamo poi New Campo Marzio, una collaborazione tra lifeforms.io e Studio Kinch, in cui si sperimentano IA e motori per *videogame* come Unreal Engine per visioni per lo spazio pubblico di Roma costruite sulla base dei disegni

di Piranesi. Il gruppo Coop Himmelblau dal 2016 ha creato il laboratorio Deep Himmelblau, dedicato alla sperimentazione dell'IA addestrata sul repertorio dei progetti dello studio. Infine, ricordiamo gli studi di Koh e la serie Neural Sampling (e gli sviluppi 3D-GAN-Ar-chair-tecture del 2020 e 3D-Autoregressive-Chair del 2021), dove elabora un campionamento di forme 3D per mezzo di reti neurali 3D-GAN addestrate su migliaia di modelli geometrici di sedie ed edifici a differenti scale e risoluzioni [Koh 2022]. I risultati, nella loro indeterminatezza e assenza di riferimenti stilistici e culturali rappresentano una sfida diretta all'approccio *form-finding* che mira ad arrivare a una forma ideale sulla base di vincoli e parametri fisici.

Metodi e tecniche di elaborazione: *exempla*

Le rappresentazioni realizzate nel corso della ricerca hanno consentito di approfondire e sperimentare alcuni fra i principali processi elaborativi oggi offerti dai software che utilizzano le potenzialità dell'IA. In particolare, sono state sperimentate 3 metodiche elaborative: quella offerta dai sistemi di generazione *text-to-image*; quella della modellazione che possiamo in generale definire 'morfo-generativa'; infine sono state verificate delle interazioni tra le due. Come già anticipato, tali metodiche sono state strutturate rispetto a due principali linee di ricerca visiva: *transizioni semantiche di immagini piane* (figg. 0-9); *transizioni di conformazioni spaziali tridimensionali* (fig. 10). Per sperimentare i vari processi di generazione di immagini bidimensionali, abbiamo predisposto un sistema di lavoro incentrato soprattutto sul software Midjourney. La prima operazione in questi sistemi elaborativi di IA è la formulazione delle domande che poniamo alla macchina che, come prevedibile, diventa un'azione fondamentale. È nella scelta delle parole chiave contenute nella domanda che possiamo destabilizzare la risposta, in parte prevederla ma, come abbiamo più volte verificato, anche stupirci delle risposte. Le prime sperimentazioni effettuate sono state svolte sul tema della generazione di uno 'stile visivo' riferito all'interrogazione. Questa prima sperimentazione è stata chiamata 'Esercizi di Stile, monotematici autoriali', sia attraverso il testo, sia utilizzando immagini (figg. 1-3). Abbiamo così iniziato a testare il software sul concetto di stile riferito all'opera di un autore specifico, attraverso parole chiave inserite nella riga di comando testuale. Dopo aver immesso il nome dell'autore studiato, sono stati inseriti anche dei riferimenti progettuali e 'tipologici', che ci hanno consentito di implementare la rispondenza dell'immagine di output. In alcuni casi, non bastando solo le parole chiave, sono state inserite alcune immagini inerenti opere specifiche del progettista, verificando così risposte visive più esaustive e appropriate. La sperimentazione è proseguita andando ad effettuare 'Ibridazioni architettoniche, autoriali e tipologiche' (figg. 4-6). L'Input dato alla macchina prevede sia l'inserimento di un prompt testuale, composto da autore, stile di rappresentazione, tipologia architettonica e metodo di rappresentazione, sia da elaborati grafici inseriti per fornire specifici dettagli riguardo la composizione dell'immagine di output. La terza parte elaborativa sviluppata con Midjourney è stata quella denominata 'Morfemi, coppie duali e coppie oppositive' (figg. 7-9). Lo schema di partenza è quello offerto dai 'Morfemi' elaborati da Franco Purini, e in questa direzione dopo aver inserito alcuni riferimenti testuali, abbiamo inserito due immagini di riferimento; il software sembra così aver riconosciuto il 'linguaggio progettuale', riproducendone il senso (fig. 7. Tav. 3A 1). Discussi i risultati, abbiamo provato a generare delle forme tridimensionali reinserendo l'immagine generata in precedenza e digitando nella riga di comando testuale '*Axonometric Projections*'. I risultati sono stati interessanti, soprattutto perché il software, oltre ad aver riconosciuto il metodo di rappresentazione, ha conformato credibili interpretazioni tridimensionali del progetto studiato (figg. 7. Tav. 3A 2-5). Altri interessanti risultati sono pervenuti dalla digitazione di parole chiave nell'ambito dello studio di 'coppie duali' e 'coppie oppositive'. Infatti, dopo aver delineato gli argomenti da sperimentare, come ad es. il concetto architettonico di 'pieno e vuoto' (fig. 8) e di 'fluido e rigido' (fig. 9), abbiamo iniziato la sperimentazione inserendo esclusivamente, nella riga di comando, input testuali, senza il supporto di immagini di riferimento. Dalla valutazione dell'esperienza svolta, osservando alcune potenzialità di ordine morfo-generativo del processo elaborativo utilizzato (cfr. fig. 7. Tav. 3A 2-5), si è voluto andare oltre e sperimentare delle transizioni di conformazioni di

modelli tridimensionali digitali. Si è così proceduto alla modellazione (utilizzando soprattutto software di grafica vettoriale 3D, in particolare 3ds Max) di semplici forme tridimensionali poi trasformate in sequenze attraverso dei semplici modificatori (ad es. *particolar system*, *metaball*, *blob*, ecc.). Verificato l'esito si è successivamente intervenuti a livello formale per tracciare delle trasformazioni più coerenti con il progetto che ci eravamo prefissati. A questo punto, raggiunto un sufficiente livello trasformativo (fig.10. Tav. 5.1), abbiamo implementato il processo elaborativo attraverso le potenzialità di Midjourney. Dopo aver inserito delle parole chiave come, ad es., '*Axonometric projections, Isometric projections, transitions, trasformations*'. Il risultato è stato molto interessante, soprattutto a livello creativo (semantico), avvicinando questo processo elaborativo alle tematiche del parametrico (fig. 10. Tav. 5.2-5.6)

Note

[1] Il contributo è stato articolato rispetto a tre ambiti tematici, ognuno trattato individualmente da un autore.
Il primo tema, 'Modelli semantici in cerca di autore, fra *Soft Machine* e *Hard Machine*' è stato elaborato da Maurizio Unali, e introduce alcuni aspetti fondativi inerenti la struttura dell'intera ricerca, analizzando i principali risultati raggiunti e evidenziando alcuni possibili sviluppi.
Il secondo tema, 'Dal riconoscimento di immagini al *text-to-image* fino *text-to-3D*' è stato elaborato da Giovanni Caffio, e partendo da un excursus sullo sviluppo delle tecnologie computazionali arriva poi ad evidenziare alcune fra le più interessanti progettualità elaborate nell'ambito dei rapporti fra la rappresentazione architettonica e IA.
Infine, il contributo di Fabio Zollo, dal titolo 'Metodi e tecniche di elaborazione: *exempla*', illustra le principali tecnoculture utilizzate per elaborare le immagini e gli spazi realizzati nel corso della ricerca.

[2] <www.lineamenta.it>, <www.instagram.com/atlasofvirtualcity/> Rappresentazioni che espandono i propri confini disciplinari, scambiando i propri geni in un mutare continuo, elaborando processi metamorfici, fino a configurare un nuovo modello spazio-temporale abitabile virtualmente, in cui è ridotto il divide tra analogico e digitale. Si evidenziano così alcuni possibili sviluppi della ricerca; una domanda aperta che attualizza soprattutto lo storico tema dell'autorialità: siamo ancora in grado di identificare creatività (testi, immagini, musica, ecc.) generate dalle intelligenze artificiali?.

[3] <https://aiartists.org/alexander-mordvintsev>.

[4] <https://openai.com/blog/clip/>.

[5] <https://openai.com/dall-e-2/>.

[6] <https://midjourney.com/>.

[7] <https://stablediffusionweb.com/>.

[8] <https://vimeo.com/552788561>.

[9] <https://www.isomorphism-jeanjaminet.com/The-Serlio-Code>.

Riferimenti bibliografici

Caffio G., Unali M. (2022). Verso una storia dell'Abitare Virtuale. Dal Cyberspace a Second Life fino al Metaverso di Facebook e oltre. In C. Battini, E. Bistagnino (a cura di). *Dialoghi. Visioni e visualità. Testimoniare Comunicare Sperimentare. Atti del 43° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione*, pp. 205-220. Milano: FrancoAngeli.

Del Campo M. (2022). *Neural Architecture*. {*Design and Artificial Intelligence*}. San Francisco: Applied Research & Design.

Del Campo M., Leach N. (Eds.). (2022). *Machine Hallucinations. Architecture and Artificial Intelligence*. Architectural Design, Vol. 92(3). Hoboken: John Wiley & Sons.

Esser P., Rombach R., Ommer O. (2021). *Taming Transformers for High-Resolution Image Synthesis (a.k.a #VQGAN)*. ArXiv.

Gatys L.A., Ecker A.S., Bethge M. (2015). A Neural Algorithm of Artistic Style. In *ArXiv*.

Goodfellow I. et al. (2014). Generative Adversarial Networks. In *ArXiv*.

Karras T., Laine S., Aila T. (2018). A Style-Based Generator Architecture for Generative Adversarial Networks. In *ArXiv*.

Koh I. (2022). Architectural Plasticity: The Aesthetics of Neural Sampling. In M. Del Campo, N. Leach (Eds.). *Neural Architecture*. {*Design and Artificial Intelligence*}, pp. 86-93. San Francisco: Applied Research & Design.

Mordvintsev A., Olah, C., Tyka M. (1 luglio 2015). DeepDream – a code example for visualizing Neural Networks. <https://ai.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html> (consultato il 6 febbraio 2023).

Poole B., Jain A., Barron J.T., Mildenhall B. (2022). DreamFusion: Text-to-3D using 2D Diffusion. In *ArXiv*. <https://dreamfusion3d.github.io/>(consultato il 2 febbraio 2023).

Rayner A. (28 marzo 2016). Can Google's Deep Dream become an art machine? *The Guardian*. <https://www.theguardian.com/artanddesign/2016/mar/28/google-deep-dream-art> (consultato il 6 febbraio 2023).

Rodari G. (2010). *Grammatica della fantasia*. San Dorligo della Valle (TS): Einaudi Ragazzi.

Sacchi L. (2021). *Il mestiere di architetto*. Siracusa: LetteraVentidue.

Sacchi L. (2022). Intelligenza artificiale e architettura. In *Op.cit. Selezione della critica d'arte contemporanea,* vol. 174, n. 5, pp. 5-15.

Unali M. (2001). *Pixel di architettura.* Roma: Edizioni Kappa.

Unali M. (2018). Note sul ruolo rappresentativo/conformativo del disegno di architettura contemporaneo. In C.L. Marcos et al. (a cura di). *De trazos, huellas e improntas. Arquitectura, ideación, representación y difusión,* Proceedings of the XVII International Conference EGA 2018, Alicante 30 maggio - 1 giugno 2018, pp. 109-116. Alicante: Universidad de Alicante.

Unali M. (2019). Architettura e tecnocultura 'post' digitale. Verso una storia. In *Op.cit. Selezione della critica d'arte contemporanea*, n. 164, pp. 5-21.

Unali M., Caffio G. (2023). Semantic Model Learning. In D. Villa, F. Zuccoli (a cura di). *Proceedings of the 3rd International and Interdisciplinary Conference on Image and Imagination. IMG 2021*. Lecture Notes in Networks and Systems, vol. 631. Cham: Springer.

Wiggers K. (20 dicembre 2022). *OpenAI releases Point-E, an AI that generates 3D models. TechCrunch+.* <https://techcrunch.com/2022/12/20/openai-releases-point-e-an-ai-that-generates-3d-models/> (consultato il 6 febbraio 2023).

Zhu J., Park T., Isola P., Efros A.A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In *ArXiv*.

Autori
Maurizio Unali, Università degli Studi "G. d'Annunzio" Chieti-Pescara, maurizio.unali@unich.it
Giovanni Caffio, Università degli Studi "G. d'Annunzio" Chieti-Pescara, giovanni.caffio@unich.it
Fabio Zollo, Università degli Studi "G. d'Annunzio" Chieti-Pescara, fabio.zollo@unich.it

Per citare questo capitolo: Unali Maurizio, Caffio Giovanni, Zollo Fabio (2023). Transizioni d'immagini e architetture al tempo dell'IA. Modelli semantici in cerca di autore/Transitions of Images and Architectures in the Time of AI. Semantic Models in Search of an Author. In Cannella M., Garozzo A., Morena S. (a cura di). *Transizioni. Atti del 44° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Transitions. Proceedings of the 44th International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 3226-3243.

 Isbn 9788835155119

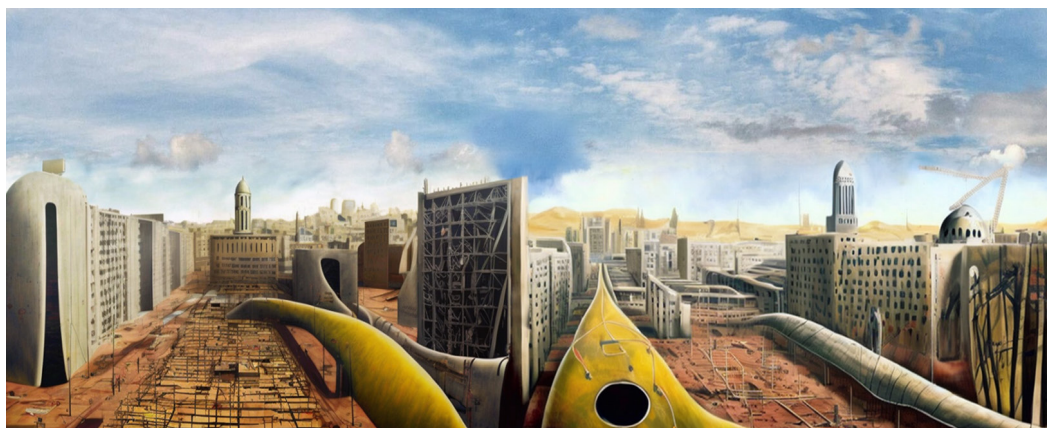
44th International Conference
of Representation Disciplines Teachers
Congress of Unione Italiana per il Disegno

Transitions of Images and Architectures in the Time of AI. Semantic Models in Search of an Author

Maurizio Unali
Giovanni Caffio
Fabio Zollo

Abstract

This essay summarizes the latest outcome of a long research path focused on the study of the representational and conformative role of the semantic model in architecture, particularly in light of recent technocultural innovations introduced by the concept of Artificial Intelligence. The idea of a semantic model has been verified and tested through the most recent elaborative processes of representation, which use AI in various ways, from text-to-image technology to morpho-generative parametric modeling and beyond. These post-digital designs enable new transitions of meaning fueled by cultural hybridizations and aesthetic recycling found in every disciplinary field, but particularly in architectural representation, which provides a fruitful laboratory of experimentation for configuring the forms of virtual living, even in the dimension of the metaverse.
The research was structured in two main phases. In the first phase, after studying the main antecedents and conceptual and visual references, some of the many designs elaborated on the theme by architects, designers and artists were catalogued and classified, highlighting semantics, methods and techniques of elaboration. In the second phase thematic designs were prepared that, between research and teaching, allowed us to elaborate 'transitions' of images and spaces through technologies referable in general to AI. In conclusion, the processed images were published and the methodologies adopted were illustrated.

Keywords
Artificial Intelligence, Semantic Models, Text-to-Image, Representative/Conformative, Virtual Living

Fantasy. *Analogous city in the age of AI*, 2023. Image generated with Midjourney, prompt by Authors, text: 'Arduino Cantafora, Analogous City, Costa, Gianni Rodari, Grammar of Fantasy'.

doi.org/10.3280/oa-1016-c464

Introduction

The research presented here is the latest result of a long study (still ongoing) focused on the elaboration of the representational and conforming role [Unali 2018] of drawing for the 'semantic model' in architecture and design [Unali, Caffio 2023], here updated with respect to the latest technocultural innovations, introduced above all by the concept of Artificial Intelligence, with the consequent interdisciplinary design.
The research was structured in two phases of development, summarised here in three chapters [1]. In the first phase – *Survey of the State of the Art* – after studying the main background and conceptual and visual references that the topic draws from the history of representation (and beyond), some of the many designs elaborated on the topic by architects, designers and artists were catalogued and classified, highlighting semantics, methods and techniques of elaboration. In the second phase of the research – *Experimentations* –, after initial applicative tests mainly of aesthetic (semantic models) and technological (methods and techniques) – fed mainly by the knowledge acquired in the previous phase –, thematic designs were prepared with respect to a more mature structure that, between research and teaching, allowed us to elaborate conscious 'transitions' of images and spaces through technologies referable in general to AI, thus addressing the various aspects opened up by the topic.

Semantic models in search of an author, between 'Soft Machine' and 'Hard Machine'

Since the early days of the computer revolution, every expressive language has wondered about the complex and fascinating relationships between our human 'Soft Machine' – quoting the title of William S. Burroughs' novel – and the artificial 'Hard Machine' of the digital (such as HAL 9000 in Stanley Kubrick's *2001: A Space Odyssey*). The examples are endless, from literature to film, science to art, philosophy and beyond.
In the first phase of the research carried out – *Survey of the State of the Art* – the study of the references that have in more ways than one fueled the thought and action of architectural design in the creative dimension today attributable to AI processes [Sacchi 2022], were fundamental and numerous.
Among them, we mention here, for reasons of space, only very few background and conceptual references to recall that precious red thread contained in the history of representation: Walter Benjamin's *The Work of Art in the Age of its Technical Reproducibility* (1936); 'The Medium is the Message', anticipated since '67 by Marshall McLuhan; Michelangelo Pistoletto's virtual *Mirroring Paintings* (from the 1960s); the conceptual synthesis proposed by Peter Eisenman for the *Virtual House* (1997), among the first examples of a conformative-creative use of modeling processes; the concept of *Big Data* studied in '99 by researchers Bryson, Kenwright, Cox, Ellsworth and Haimes [Sacchi 2021]; *AI, Artificial Intelligence* (2001) by Steven Spielberg, which opens the topic to the general public; ChatGPT (Chat Generative Pre-trained Transformer), an artificial intelligence-based chatbot created by OpenAI in 2022, ; the auction sale at Christie's in 2018 of the 'Portrait of Edmond Belamy', a work of art generated with AI by the Obvious Collective. The knowledge gained is the basis for the second phase of the research - Experimentations -, where thematic representations were realised through the use of AI (from text-to-image to morpho-generative modelling) with respect to a working structure that allowed us to elaborate conscious 'transitions' of images and spaces.
The processed representations were then structured with respect to two main lines of research identified in our AI techno-cultural experiments: semantic transitions of plane images; transitions of three-dimensional spatial conformations. Within the first group, mainly using the Midjourney platform, four macro-themes of experimentation were designed: 1. Style Exercises (figs. 1-3); 2. Architectural Hybridizations (figs. 4-6); 3. Morphemes, dual pairs and oppositional pairs (figs. 7-9); 4. Fantasy (cover image). In the second group, after the experience gained from the previous elaborative processes, morphological transformations of 3D spaces were experimented with (fig. 10).

After the analysis of the conformative results achieved by the research, the great theoretical (semantic) and visual (technocultural) potential that AI today offers to research and didactics of representation is confirmed. Shared creativities between 'Soft and Hard Machine' (greatly simplifying), which generate semantic models 'in search of an author' – as 'participatory design' –, the outcome of a new elaborative process in which cultural processes, ideas and images preparatory to the design of new forms of living converge and intersect (in the awareness of historical references). Designs that enable new transitions of meaning fueled by the cultural hybridizations and aesthetic recycling found in every disciplinary field, but that in architectural representation find a fruitful laboratory of experimentation to configure the forms of contemporary virtual living [Caffio, Unali 2022], even in the current dimension of the postdigital [Unali 2019] metaverse [2].

Fig. 1. Tab. 1A (1-5). Exercises in Style, single-issue authorship, 2023. Image generated with Midjourney, requested by the Authors, text: 'Le Corbusier's Architecture'.

Fig. 2. Tab. 1B (1-5). Exercises in Style, single-issue authorship, 2023. Image generated with Midjourney, requested by the Authors, text: 'Architecture of Louis Kahn'.

Fig. 3. Tab. 1C (1-5). Exercises in Style, single-issue authorship, 2023. Image generated with Midjourney, prompt by Authors, text: 'Aldo Rossi Architecture'.

Fig. 4. Tab. 2A (1-5). Architectural, authorial and typological hybridities, 2023. Image generated with Midjourney, prompt by Authors, text: 'Le Corbusier Architecture, Imaginary Prisons Giovanni Battista Piranesi'.

Fig. 5. Tab. 2B (1-5). Architectural, authorial and typological hybridities, 2023. Image generated with Midjourney, prompt by Authors, text: 'Imaginary Prisons Giovanni Battista Piranesi, La Muralla Roja by Bofil, painting Escher style'.

Fig. 6. Tab. 3B (1-5). Architectural, authorial and typological hybridities, 2023. Image generated with Midjourney, prompt by Authors, text: 'Louis Kahn, Aldo Rossi Teatro del Mondo'.

From image recognition to text-to-image and to text-to-3D

It was in 2015 when Google engineer Alexander Mordvintsev discovered that a neural network, properly trained, may be able not only to recognize and classify images but can reverse the process by creating new images from the initial ones, in actual fact paving the way for cre-

ative and 'artistic' use of AI. In Mordvintsev's words, "The experiment was to try to create a network that added details to a real image to effect an increase in its resolution. It happened that it added some details, but not the ones I expected" [3].
This seemingly fortuitous discovery led to the development of DeepDream [Mordvintsev, Olah, Tyka 2015], the first software capable of creating images with an 'artistic' flavor, characterized by a dreamlike, hallucinatory tone that intrigued and generated extensive debates within art critics [Rayner 2016]. These new possibilities have since been further developed to the point of being able to transfer stylistic or morphological features from one context to another or to implement special skills in the analysis and reproduction of selected sets of images such as may be the works produced by a single artist [Gatys, Ecker, Bethge 2015].

Fig. 7. Tab. 3A (1-5). Morphemes and Dual Pairs, 2023. Image generated with Midjourney, prompt by Authors, text: 'Franco Purini, Morfemi, Axonometric Projections'.

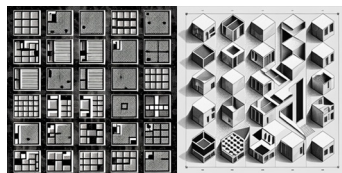



Fig. 8. Tab. 3B (1-5). Morphemes and Dual Pairs, 2023. Image generated with Midjourney, prompt by Authors, text: 'Concept of Full and Empty in Architecture'.

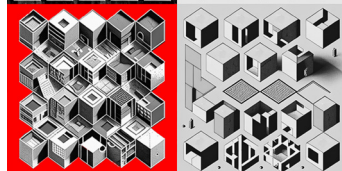
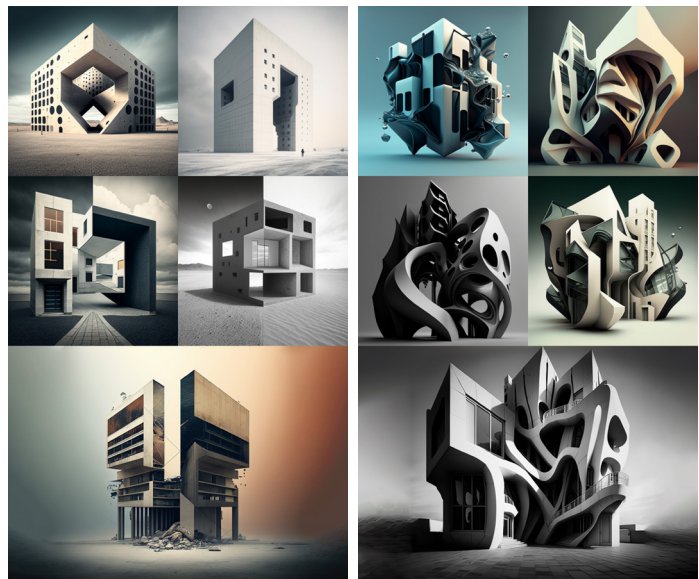
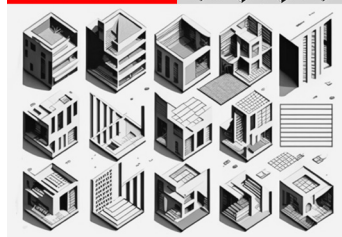



Fig. 9. Tab. 3C (1-5). Morphemes and Dual Pairs, 2023. Image generated with Midjourney, prompt by Authors, text: 'Fluid Architecture combinated with Rigid (Solid) Architecture'.

However, the most significant turning point is related to the invention of Generative Adversial Networks (GANs) by computer scientist Ian GoodFellow [Goodfellow et al. 2014]. These neural networks are based on a pair of networks that operate in competition with each other. One is tasked with generating images, the other with comparing them against a reference and training dataset. Thus, a true internal dialectic, made up of thousands of trials and errors, develops between one system that tirelessly produces images and another that discards them until they do not match the predetermined judgment criteria. This process is intended to train the image-generating network and, once achieved, while the discriminating network stops, the generating network continues to produce new representations. There are several forms of GAN. One of the most common is StyleGAN [Karras, Laine, Aila 2018], which uses the interpolation procedure to generate new variations from a repertoire of thousands of images. Other GAN networks, such as CycleGAN [Zhu, Park, Isola, Efros 2017], instead of a single dataset, use two dichotomous sets working no longer through interpolation but by extrapolation. In this way, the images produced appear to be the result of a more 'creative' process as they introduce different and innovative elements than those that can be derived exclusively from a single set.
Skipping a few steps, it is interesting to think that such elaborative processes have always been present in human beings who, through the digital expands their imaginative potential. In this regard, we recall the experience in the early 1970s of Gianni Rodari who introduced his method of the 'fantastic binomial': "A story can only be born from a 'fantastic binomial'. [...] A certain distance between the two words is needed, it is necessary that one be sufficiently foreign to the other, and their juxtaposition discreetly unusual, for the imagination to be forced to set in motion to establish a kinship between them, to build a (fantastic) whole in which the two foreign elements can coexist" [Rodari 2010, p. 16].

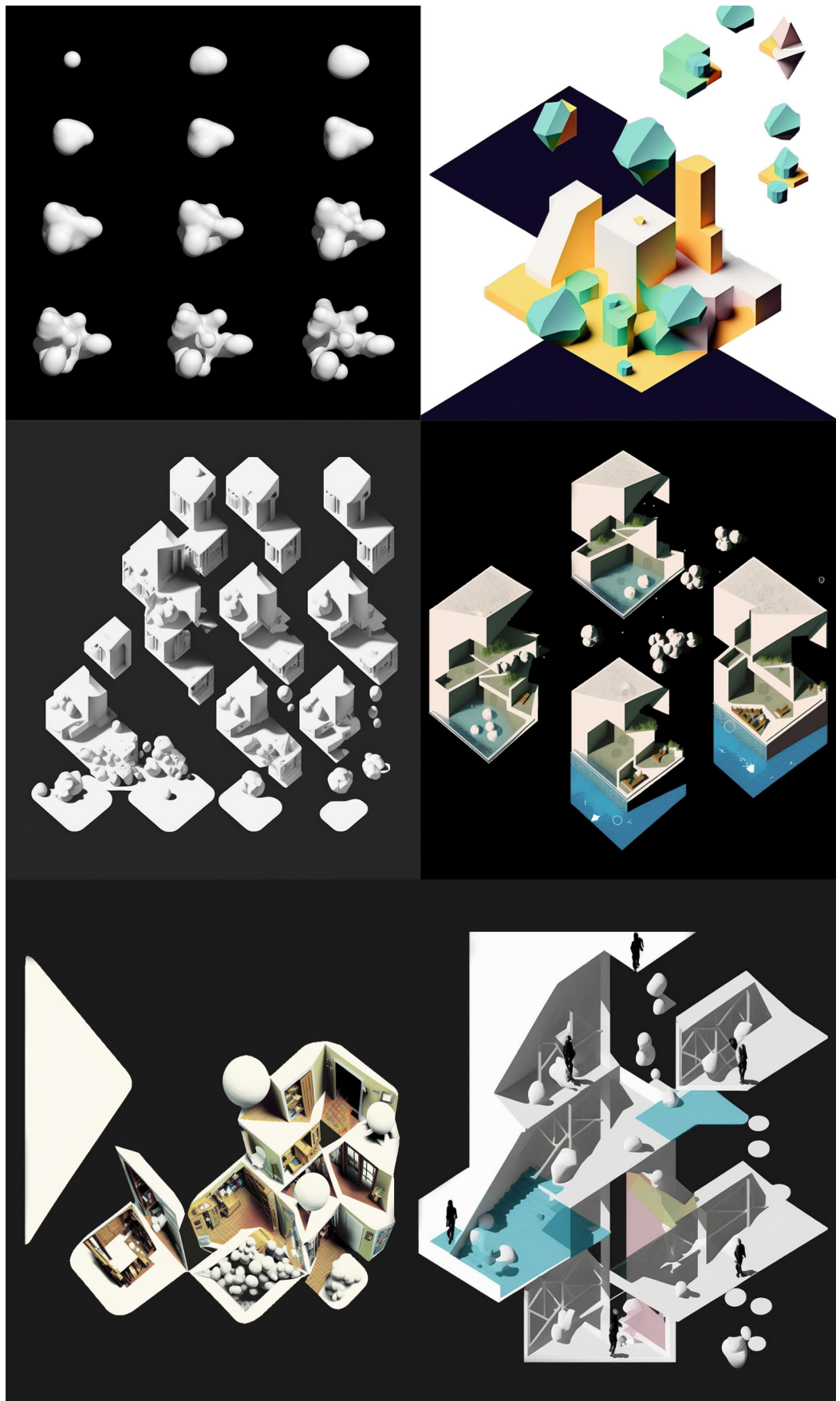



Fig. 10. Tab. 5 (1-6). Transitions of three-dimensional spatial conformations, 2023. Tab. 5.1. Morpho-generative modeling. Tabs. 5.2.-5.6. Interactions between modeler and text- to-image. Image generated with Midjourney, prompt by Authors, text: 'Axonometric Projections, Transitions'.

Returning to the recent history of AI, innovations then continued with the emergence of CLIP (Contrastive Language-Image Pre-training) [4], an image generation system that employs GPT-3 (Generative Pre-trained Transformer 3), a natural language processing method that harnesses the power of deep learning to create seemingly human-produced text and is also able to use content on the Web to associate strings of explanatory text with images.
The potential of CLIP then becomes exponential when combined with VQGans (Vector Quantized Generative Adversarial Network) [Esser, Rombach, Ommer 2021] a neural network that can combine convolutional neural networks to generate high-resolution images. The result of this union is the ability to generate two-dimensional images from textual prompts as is done with the now well-known Dall-e 2 [5], Midjourney [6] and Stable Diffusion [7].
While the generation of two-dimensional figures becomes more effective and realistic with each passing day, systems capable of translating textual prompts into three-dimensional forms are still being tested. However, due to increased computational power, improved algorithms, more competitors, and growing investment, it is not difficult to imagine substantial innovations in the near future.
Among the projects working in the direction of generating three-dimensional structures we can mention Point-e and DreamFusion. The former is a machine learning system developed by OpenAI (the same company that develops Dall-e 2 and ChatGPT) that generates three-dimensional objects based on textual prompts. Point-e does not actually create a concluded shape but rather point clouds in space that a second AI system converts into meshes [Wiggers 2022].
DreamFusion [Poole, Jain, Barron, Mildenhall 2022], on the other hand, uses a pre-trained 2D text-to-image diffusion model to achieve a 3D text-to-synthesis. The 2D diffusion model serves as the initial basis for optimizing a parametric image generator. In a procedure similar to that used in DeepDream, the 3D model is initially generated randomly (a Neural Radiance Field, or NeRF) so that its 2D representations from random angles appear sufficiently consistent. The resulting 3D model can then be viewed from any angle, illuminated by light sources, or inserted into a 3D space.
Beyond the computational advances yet to come, here we would like to highlight how there are many architects and researchers already experimenting with different approaches and solutions to the use of AIs for generating new architectural creativity. We start by mentioning the fundamental research conducted in Matias Del Campo through his texts [Del Campo 2022] and the architectural experiments conducted with his SPAN studio (with Sandra Manninger) using a standard convolutional neural network (CNN), called graphical convolutional neural network (GCNN), to model and execute the entire design process to three-dimensional models and information [Del Campo, Leach 2022, p. 45]. The Habidatum studio combines GAN neural networks and human intervention, aesthetic principles and form-finding processes to generate the configurations of the *Endless Skyscraper* series [8]. Esquivel, Bugni and Jaminet experiment with GAN networks to research 3D reinterpretations of the Serlian code in *The Serlio Code* project [9]. Then there is New Campo Marzio, a collaboration between lifeforms.io and Studio Kinch, in which they experiment with AI and video game engines such as Unreal Engine for visions for Rome's public space built on the basis of Piranesi's drawings. Since 2016, the Coop Himmelblau group has created the Deep Himmelblau laboratory, dedicated to experimenting with AI trained on the studio's repertoire of projects. Finally, let us recall Koh's studies and the *Neural Sampling* series (and the 3D-GAN-*Ar-chair-tecture* developments of 2020 and 3D-*Autoregressive-Chair* of 2021), where he processes 3D shape sampling by means of 3D-GAN neural networks trained on thousands of geometric models of chairs and buildings at different scales and resolutions [Koh 2022]. The results, in their indeterminacy and absence of stylistic and cultural references, represent a direct challenge to the form-finding approach that aims to arrive at an ideal form based on constraints and physical parameters.

Processing methods and techniques: exempla

The representations made in the course of the research allowed us to investigate and experiment with some of the main elaborative processes offered today by software using the potential of AI. In particular, 3 elaborative methods were tested: that offered by text-to-image

generation systems; that of modeling that we can generally define as 'morpho-generative'; and finally, interactions between the two were tested. As already anticipated, these methods were structured with respect to two main lines of visual research: *semantic transitions of plane images* (figs. 0-9); *transitions of three-dimensional spatial conformations* (fig. 10).
In order to experiment with the various processes of two-dimensional image generation, we set up a working system centered mainly on the Midjourney software. The first operation in these AI processing systems is the formulation of the questions we ask the machine, which, predictably, becomes a key action. It is in the choice of key words contained in the question that we can destabilize the answer, partly predict it but, as we have repeatedly verified, also be surprised by the answers. The first experiments carried out were on the topic of generating a 'visual style' referring to the interrogation. This first experimentation was called "Style Exercises, single-issue authorship," both through text and using images (figs.1-3). We thus began to test the software on the concept of style referring to the work of a specific author, through keywords entered in the text command line. After entering the name of the author studied, design and 'typological' references were also entered, which allowed us to implement the correspondence of the output image. In some cases, since keywords alone were not enough, some images pertaining to specific works of the designer were entered, thus verifying more comprehensive and appropriate visual responses.
Experimentation continued by going on to perform 'Architectural, authorial and typological hybridizations' (figs. 4-6). The Input given to the machine involves both the input of a textual prompt, consisting of author, style of representation, architectural type, and method of representation, and elaborate graphics entered to provide specific details regarding the composition of the output image. The third elaborative part developed with Midjourney was that called 'Morphemes, dual pairs and oppositional pairs' (figs. 7-9). The starting scheme is the one offered by the *Morphemes* elaborated by Franco Purini, and in this direction after entering some textual references, we inserted two reference images; thus the software seems to have recognized the 'design language', reproducing its meaning (fig. 7. Table 3A 1). Having compared the results, we tried generating three-dimensional shapes by reinserting the previously generated image and typing in the textual command line 'Axonometric Projections'. The results were interesting, especially because the software, in addition to recognizing the method of representation, conformed credible three-dimensional interpretations of the studied project (figs. 7. Tab. 3A 2-5). Other interesting results came from keyword typing as part of the study of 'dual pairs' and 'oppositional pairs'. In fact, after outlining the topics to be experimented with, e.g., the architectural concepts of 'full and empty' (fig. 8) and 'fluid and rigid' (fig. 9), we began the experimentation by entering textual imputs exclusively, in the command line, without the support of reference images. From the evaluation of the experience carried out, observing some morpho-generative potentialities of the elaborative process used (see fig. 7 Table 3A 2-5), we wanted to go further and experiment with transitions of conformations of digital three-dimensional models. Thus, modeling (mainly using 3D vector graphics software, especially 3ds Max) of simple three-dimensional shapes then transformed into sequences by means of simple modifiers (e.g., particular system, metaball, blob, etc.) was carried out. Having verified the outcome, we then intervened at the formal level to draw transformations that were more consistent with the design we had set for ourselves. At this point, having reached a sufficient transformational level (fig.10 Tab.5.1), we implemented the elaborative process through the power of Midjourney. After entering keywords such as, e.g., 'Axonometric projections, Isometric projections, transitions, transformations'. The result was very interesting, especially at the creative (semantic) level, bringing this elaborative process closer to the themes of parametric (fig. 10 Table 5.2-5.6).

Notes

[1] The contribution was articulated with respect to three thematic areas, each treated individually by an author. The first theme, 'Semantic models in search of an author, between Soft Machine and Hard Machine' was elaborated by Maurizio Unali, and introduces some founding aspects inherent to the structure of the entire research, analysing the main results achieved and highlighting some possible developments. The second topic, 'From image recognition to text-to-image up to text-to-3D' was

elaborated by Giovanni Caffio, and starting from an excursus on the development of computational technologies, it then goes on to highlight some of the most interesting projects elaborated in the field of relations between architectural representation and AI. Finally, Fabio Zollo's contribution, entitled 'Processing methods and techniques: exempla', illustrates the main technocultures used to process the images and spaces created during the research.

[2] <www.lineamenta.it>, <www.instagram.com/atlasofvirtualcity/>. Representations that expand their disciplinary boundaries, exchanging their genes in a continuous mutation, elaborating metamorphic processes, until they configure a new space-time model that can be virtually inhabited, in which the divide between analogue and digital is reduced. Some possible developments of the research are thus highlighted; an open question that especially updates the historical theme of authorship: are we still able to identify creativity (texts, images, music, etc.) generated by artificial intelligences?

[3] <https://aiartists.org/alexander-mordvintsev>.

[4] <https://openai.com/blog/clip/>.

[5] <https://openai.com/dall-e-2/>.

[6] <https://midjourney.com/>.

[7] <https://stablediffusionweb.com/>.

[8] <https://vimeo.com/552788561>.

[9] <https://www.isomorphism-jeanjaminet.com/The-Serlio-Code>.

References

Caffio G., Unali M. (2022). Toward a history of Virtual Living. From Cyberspace to Second Life to the Facebook Metaverse and beyond. In C. Battini, E. Bistagnino (Eds). *Dialogues. Visions and visuality. Witnessing Communicating Experimenting. Proceedings of the 43rd International Conference of Repre- sentation Disciplines Teachers,* pp. 205-220. Milano: FrancoAngeli.

Del Campo M. (2022). *Neural Architecture*. {*Design and Artificial Intelligence*}. San Francisco: Applied Research & Design.

Del Campo, M., Leach, N. (Eds.). (2022). *Machine Hallucinations*. *Architecture and Artificial Intelligence*. Architectural Design, Vol. 92, No. 3. Hoboken: John Wiley & Sons.

Esser P., Rombach R., Ommer O. (2021). *Taming Transformers for High-Resolution Image Synthesis (a.k.a #VQGAN).* ArXiv.

Gatys L.A., Ecker A.S., Bethge M. (2015). *A Neural Algorithm of Artistic Style*. In *ArXiv*.

Goodfellow I. et al. (2014). *Generative Adversarial Networks*. In *ArXiv*.

Karras T., Laine S., Aila T. (2018). *A Style-Based Generator Architecture for Generative Adversarial Networks*. In *ArXiv*.

Koh I. (2022). Architectural Plasticity: The Aesthetics of Neural Sampling. In M. Del Campo, N. Leach (Eds.). *Neural Architecture*. {*Design and Artificial Intelligence*}, pp. 86-93. San Francisco: Applied Research & Design.

Mordvintsev A., Olah, C., Tyka M. (1 July 2015). DeepDream – a code example for visualizing Neural Networks. <https://ai.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html> (accessed 6 February 2023).

Poole B., Jain A., Barron J.T., Mildenhall B. (2022). *DreamFusion: Text-to-3D using 2D Diffusion*. In *ArXiv*. <https://dreamfusion3d.github.io/> (accessed 6 February 2023).

Rayner A. (28 March 2016). Can Google's Deep Dream become an art machine? *The Guardian.* <https://www.theguardian.com/artanddesign/2016/mar/28/google-deep-dream-art> (accessed 6 February 2023).

Rodari G. (2010). *Grammatica della fantasia*. San Dorligo della Valle (TS): Einaudi Ragazzi.

Sacchi L. (2022). Intelligenza artificiale e architettura. In *Op.cit. Selezione della critica d'arte contemporanea,* Vol. 174, No. 5, pp. 5-15.

Sacchi L. (2021). *Il mestiere di architetto.* Syracuse: LetteraVentidue.

Unali M. (2001). *Pixel di architettura.* Rome: Edizioni Kappa.

Unali M. (2018). Note sul ruolo rappresentativo/conformativo del disegno di architettura contemporaneo. In C.L. Marcos et al. (Eds.). *De trazos, huellas e improntas. Arquitectura, ideación, representación y difusión, Proceedings of the XVII International Conference EGA 2018*, Alicante 30 May - 1 June 2018, pp. 109-116. Alicante: Universidad de Alicante.

Unali M. (2019). Architettura e tecnocultura 'post' digitale. Verso una storia. In *Op.cit. Selezione della critica d'arte contemporanea*, No. 164, pp. 5-21.

Unali M., Caffio G. (2023). Semantic Model Learning. In D. Villa, F. Zuccoli (Eds.). *Proceedings of the 3rd International and Interdisciplinary Conference on Image and Imagination. IMG 2021*. Lecture Notes in Networks and Systems, Vol. 631. Cham: Springer.

Wiggers K. (20 December 2022). OpenAI releases Point-E, an AI that generates 3D models. TechCrunch+. <https://techcrunch.com/2022/12/20/openai-releases-point-e-an-ai-that-generates-3d-models/> (accessed 6 February 2023).

Zhu J., Park T., Isola P., Efros A.A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In *ArXiv*.

Authors
Maurizio Unali, Università degli Studi "G. d'Annunzio" Chieti-Pescara, maurizio.unali@unich.it
Giovanni Caffio, Università degli Studi "G. d'Annunzio" Chieti-Pescara, giovanni.caffio@unich.it
Fabio Zollo, Università degli Studi "G. d'Annunzio" Chieti-Pescara, fabio.zollo@unich.it

To cite this chapter: Unali Maurizio, Caffio Giovanni, Zollo Fabio (2023). Transizioni d'immagini e architetture al tempo dell'IA. Modelli semantici in cerca di autore/Transitions of Images and Architectures in the Time of AI. Semantic Models in Search of an Author. In Cannella M., Garozzo A., Morena S. (Eds.). *Transizioni. Atti del 44° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Transitions. Proceedings of the 44th International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 3226-3243.

 Isbn 9788835155119